



REC'D 02 SEP 2003  
WIPO PCT

**Prioritätsbescheinigung über die Einreichung  
einer Patentanmeldung**

**Aktenzeichen:** 102 30 809.8

**Anmeldetag:** 08. Juli 2002

**Anmelder/Inhaber:** T-Mobile Deutschland GmbH, Bonn/DE

**Bezeichnung:** Verfahren zur Übertragung von Audiosignalen nach dem Verfahren der priorisierenden Pixelübertragung

**IPC:** H 03 M, G 10 L

**PRIORITY  
DOCUMENT**

SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

**Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.**

München, den 29. Juli 2003  
Deutsches Patent- und Markenamt

Der Präsident  
Im Auftrag

Steck

08.07.2002

T-Mobile Deutschland GmbH

## **Verfahren zur Übertragung von Audiosignalen nach dem Verfahren der priorisierenden Pixelübertragung**

Die Erfindung betrifft ein Verfahren zur Übertragung von Audiosignalen nach dem Verfahren der priorisierenden Pixelübertragung nach dem Oberbegriff des Patentanspruchs 1.

Zur Zeit existiert eine Vielzahl verschiedener Verfahren zur komprimierten Übertragung von Audiosignalen. Im wesentlichen existieren folgende Verfahren.

- Reduzierung der Abtastrate, z.B. 3 kHz anstelle von 44 kHz
- Nichtlineare Übertragung der Abtastwerte, z.B. bei ISDN Übertragung
- Benutzung von vorher abgespeicherten Akustiksequenzen, z.B. MIDI oder Stimmnachbildung
- Verwendung von Markov Modellen zur Korrektur von Übertragungsfehlern

Die Gemeinsamkeiten der bekannten Verfahren liegen darin, dass auch bei niedrigeren Übertragungsraten eine befriedigende Sprachverständlichkeit vorhanden ist. Dieses wird im wesentlichen durch Mittelwertbildungen erreicht. Jedoch ergeben unterschiedliche Stimmen der Quelle ähnlich klingende Stimmen in der Senke, so dass z.B. Stimmungsschwankungen, die in einem normalen Gespräch erkennbar sind, nicht mehr übertragen werden. Dadurch ergibt sich eine deutliche Einschränkung in der Kommunikationsqualität.

Verfahren zur Komprimierung und Dekomprimierung von Bild- oder Videodaten mittels priorisierter Pixelübertragung sind in den deutschen Patentanmeldungen DE 101 13 880.6 (entspricht PCT/DE02/00987) und DE 101 52 612.1 (entspricht PCT/DE02/00995) beschrieben. Bei diesen Verfahren werden z.B. digitale Bild- oder

Videodaten bearbeitet, die aus einem Array einzelner Bildpunkte (Pixel) bestehen, wobei jedes Pixel einen sich zeitlich verändernden Pixelwert aufweist, der Farb- oder Helligkeitsinformation des Pixels beschreibt. Erfindungsgemäß wird jedem Pixel bzw. jeder Pixelgruppe eine Priorität zugeordnet und die Pixel entsprechend ihrer Priorisierung in einem Prioritätenarray abgelegt. Dieses Array enthält zu jedem Zeitpunkt, die nach der Priorisierung sortierten Pixelwerte. Entsprechend der Priorisierung werden diese Pixel, und die für die Berechnung der Priorisierung benutzten Pixelwerte, übertragen bzw. abgespeichert. Ein Pixel bekommt eine hohe Priorität, wenn die Unterschiede zu seinen benachbarten Pixel sehr groß sind. Zur Rekonstruktion werden die jeweils aktuellen Pixelwerte auf dem Display dargestellt. Die noch nicht übertragenden Pixel werden aus den schon übertragenden Pixel berechnet. Diese Verfahren lassen sich prinzipiell auch für eine Übertragung von Audiosignalen verwenden.

Die Aufgabe der Erfindung besteht deshalb darin, ein Verfahren zur Übertragung von Audiosignalen anzugeben, das auch bei niedrigen Übertragungsbandbreiten möglichst verlustfrei arbeitet.

Diese Aufgabe wird erfindungsgemäß durch die Merkmale des Patentanspruchs 1 gelöst.

Gemäß der Erfindung wird das Audiosignal zunächst in eine Anzahl  $n$  von spektralen Anteilen zerlegt. Das zerlegte Audiosignals wird in einem zweidimensionalen Array mit einer Vielzahl von Feldern gespeichert, mit Frequenz und Zeit als Dimensionen und der Amplitude als jeweils einzutragenden Wert im Feld. Dann werden aus jedem einzelnen Feld und mindestens zwei zu diesem Feld benachbarten Feldern des Arrays Gruppen gebildet, und den einzelnen Gruppen eine Priorität zugeordnet, wobei die Priorität einer Gruppe umso größer gewählt wird, je größer die Amplituden der Gruppenwerte sind und/oder je größer die Amplitudenunterschiede der Werte einer Gruppe sind und/oder je näher die Gruppe an der aktuellen Zeit liegt.

Schließlich werden die Gruppen in der Reihenfolge ihrer Priorität an den Empfänger übertragen.

Das neue Verfahren beruht im wesentlichen auf den Grundlagen von Shannon. Demnach lassen sich Signale verlustfrei übertragen, wenn man sie mit der doppelten Frequenz abtastet. Das bedeutet, dass der Schall in einzelne Sinusschwingung unterschiedlicher Amplitude und Frequenz zerlegbar ist. Demnach lassen sich akustische Signale eindeutig durch Übertragung der einzelnen Frequenzanteile, inklusive der Amplituden und Phasen, ohne Verluste wieder herstellen. Hierbei wird auch insbesondere ausgenutzt, dass die häufig vorkommenden Schallquellen, z.B. Musikinstrumente, menschliche Stimme, aus Resonanzkörpern bestehen, deren Resonanzfrequenz sich nicht bzw. nur langsam ändert.

Vorteilhafte Ausgestaltungen und Weiterbildungen der Erfindung sind in den abhängigen Patentansprüchen angegeben.

Ein Ausführungsbeispiel der Erfindung wird nachfolgen beschrieben. Hierbei sei insbesondere auch auf die Beschreibung und die Zeichnungen der älteren Patentanmeldungen DE 101 13 880.6 und DE 101 52 612.1 Bezug genommen.

Zunächst wird der Schall aufgenommen, in elektrische Signale umgewandelt und in seine Frequenzanteile zerlegt. Dieses kann entweder durch FFT (Fast-Fourier Transformation) oder durch n-einzelne frequenzselektierende Filter geschehen. Werden n-einzelne Filter verwendet, so nimmt jeder Filter nur eine einzelne Frequenz, bzw. ein schmales Frequenzband, auf (ähnlich den Härchen im menschlichen Ohr). Somit hat man zu jedem Zeitpunkt die Frequenz, und den Amplitudenwert bei dieser Frequenz. Dabei kann die Zahl n entsprechend der Endgeräteeigenschaften unterschiedliche Werte annehmen. Je größer n ist, desto besser kann das Audiosignal reproduziert werden. Somit ist n ein Parameter mit dem die Qualität der Audioübertragung skaliert werden kann.

Die Amplitudenwerte werden in den Feldern eines 2-dimenionalen Arrays zwischengespeichert.

Dabei entspricht die erste Dimension des Arrays der Zeitachse und die zweite Dimension der Frequenz. Damit ist jeder Abtastwert mit jeweiliger Amplitudenwert und Phase eindeutig bestimmt und kann im zugeordneten Feld des Arrays als Imaginäre Zahl abgespeichert werden. Das Sprachsignal wird somit in drei akustischen Dimensionen (Parametern) im Array dargestellt: Die Zeit z.B. in Millisekunden (ms), perzeptiv als Dauer wahrgenommen, als die erste Dimension des Arrays, die Frequenz in Hertz (Hz), perzeptiv als Tonhöhe wahrgenommen, als die zweite Dimension des Arrays und die Energie (bzw. Intensität) des Signals, perzeptiv als Lautstärke bzw. Intensität wahrgenommen, welche als Zahlenwert im entsprechenden Feld des Arrays gespeichert wird.

Im Vergleich zu den Anmeldungen DE 101 13 880.6 und DE 101 52 612.1 entspricht z.B. die Frequenz der Bildhöhe, die Zeit der Bildbreite und die Amplitude des Audiosignals (Intensität) dem Farbwert.

Ähnlich dem Verfahren der Priorisierung von Pixelgruppen bei der Bild/Videokodierung werden aus benachbarten Werten Gruppen gebildet und diese priorisiert. Jedes Feld für sich betrachtet bildet zusammen mit mindestens einem, vorzugsweise jedoch mehreren benachbarten Feldern eine Gruppe. Die Gruppen bestehen aus dem Positionswert, definiert durch Zeit und Frequenz, dem Amplitudenwert am Positionswert, und die Amplitudenwerte der umliegenden Werte entsprechend einer vorher festgelegten Form (siehe Figur 2 der Anmeldungen DE 101 13 880.6 und DE 101 52 612.1). Dabei bekommen insbesondere diejenigen Gruppen eine sehr hohe Priorität, die nahe der aktuellen Zeit liegen, und/oder deren Amplitudenwerte im Vergleich zu den anderen Gruppen sehr groß sind und/oder bei denen sich die Amplitudenwerte innerhalb der Gruppe stark voneinander unterscheiden. Die Pixelgruppenwerte werden absteigend sortiert und in dieser Reihenfolge gespeichert bzw. übertragen.

Die Breite des Arrays (Zeitachse) besitzt vorzugsweise nur eine begrenzte Ausdehnung (z.B. 5 Sekunden), d.h. es werden immer nur Signalabschnitte von z.B.

5 Sekunden Länge verarbeitet. Nach dieser Zeit (z.B. 5 Sekunden) wird das Array mit den Werten des nachfolgenden Signalabschnitts gefüllt.

Entsprechend der oben beschriebenen Priorisierungsparameter (Amplitude, zeitnahe Position und Amplitudenunterschiede zu benachbarten Werten) werden die Werte der einzelnen Gruppen im Empfänger empfangen.

Beim Empfänger werden die Gruppen wieder in ein entsprechendes Array eingetragen. Entsprechend der Patentanmeldungen DE 101 13 880.6 und DE 101 52 612.1 kann dann aus den übertragenden Gruppen wieder die dreidimensionale Spektraldarstellung erzeugt werden. Je mehr Gruppen empfangen wurden, umso genauer wird die Rekonstruktion. Die noch nicht übertragenen Arraywerte werden mittels Interpolation aus den schon übertragenden Arraywerten berechnet. Aus dem so erzeugten Array wird dann im Empfänger ein entsprechendes Audiosignal generiert, welches dann in Schall umgewandelt werden kann.

Zur Synthese des Audiosignals können z.B. n Frequenzgeneratoren verwenden werden, deren Signale zu einem Ausgangssignal addiert werden. Durch diesen parallelen Aufbau von n Generatoren ist eine gute Skalierbarkeit gegeben. Zudem kann die Taktrate durch parallele Verarbeitung drastisch reduziert werden, so dass durch einen geringeren Energieverbrauch die Wiedergabezeit bei mobilen Endgeräten erhöht wird. Für den parallelen Einsatz könnten z.B. FPGA's oder ASIC's einfacher Bauart benutzt werden.

Das beschriebene Verfahren ist nicht auf Audiosignale beschränkt. Das Verfahren kann insbesondere überall dort effektiv zur Anwendung kommen, wo mehrere Sensoren (Schallsensoren, Lichtsensoren, Tastsensoren, usw.) verwendet werden, die kontinuierlich Signale messen, die dann in einem Array (n-ter Ordnung) dargestellt werden können.

Die Vorteile gegenüber bisherigen Systemen liegen in der flexiblen Einsetzbarkeit bei erhöhten Kompressionsraten. Durch Benutzung eines Arrays, welches aus unterschiedlichen Quellen gespeist wird, erhält man automatisch eine

Synchronisation der unterschiedlichen Quellen. Eine entsprechende Synchronisation muss bei herkömmlichen Verfahren durch besondere Protokolle, bzw. Maßnahmen gesichert werden. Insbesondere bei Videoübertragung mit großen Laufzeiten, z.B. Satellitenverbindungen, wo Ton und Bild über verschiedene Kanäle übertragen werden, fällt häufig eine fehlende Synchronisation der Lippen zu der Sprache auf. So etwas kann durch das beschriebene Verfahren beseitigt werden

Da das gleiche Grundprinzip der priorisierenden Pixelgruppenübertragung sowohl bei Sprache, Bild und Videoübertragung genutzt werden kann, ist ein starker Synergieeffekt bei der Implementierung nutzbar. Außerdem kann auf diese Weise eine einfache Synchronisation zwischen Sprache und Bildern erfolgen. Außerdem könnte beliebig zwischen Bild- und Audioauflösung skaliert werden.

Betrachtet man eine einzelne Audioübertragung nach dem neuen Verfahren, so ergibt sich bei Sprache eine natürlichere Wiedergabe, da die für jeden Menschen typischen Frequenzanteile (-gruppen) mit höchster Priorität und damit verlustfrei übertragen werden.

## Patentansprüche

1. Verfahren zur Übertragung von Audiosignalen zwischen einem Sender und mindestens einem Empfänger nach dem Verfahren der priorisierenden Pixelübertragung, gekennzeichnet durch die Schritte:
  - a) Zerlegen des Audiosignals in eine Anzahl  $n$  von spektralen Anteilen,
  - b) Speichern des zerlegten Audiosignals in einem zweidimensionalen Array mit einer Vielzahl von Feldern, mit Frequenz und Zeit als Dimensionen und der Amplitude als jeweils einzutragenden Wert im Feld,
  - c) Bilden von Gruppen aus jedem einzelnen Feld und mindestens zwei zu diesem Feld benachbarten Feldern des Arrays,
  - d) Zuordnen einer Priorität zu den einzelnen Gruppen, wobei die Priorität einer Gruppe umso größer wird, je größer die Amplituden der Gruppenwerte sind und/oder je größer die Amplitudenunterschiede der Werte einer Gruppe sind und/oder je näher die Gruppe an der aktuellen Zeit liegt, und
  - e) Übertragen der Gruppen in der Reihenfolge ihrer Priorität an den Empfänger.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass das gesamte Audiosignal als Audiodatei vorliegt und als Ganzes bearbeitet und übertragen wird.
3. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass nur jeweils ein Teil des Audiosignals bearbeitet und übertragen wird.
4. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass das Audiosignal mittels FFT in seine spektralen Anteile zerlegt wird.

5. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass das Audiosignal durch eine Anzahl  $n$  von frequenzselektierenden Filtern in seine spektralen Anteile zerlegt wird.
6. Verfahren nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, dass im Empfänger die nach ihrer Priorität übertragenen Gruppen einem entsprechenden Array zugeordnet, wobei die noch nicht übertragenen Werte des Arrays aus den bereits vorhandenen Werten durch Interpolation berechnet werden.
7. Verfahren nach einem der Ansprüche 1 bis 6, dadurch gekennzeichnet, dass aus den im Empfänger vorliegenden und berechneten Werten ein elektrisches Signal generiert und in ein Audiosignal umgewandelt wird.

## **Zusammenfassung**

Die Erfindung betrifft ein Verfahren zur Übertragung von Audiosignalen zwischen einem Sender und mindestens einem Empfänger nach dem Verfahren der priorisierenden Pixelübertragung. Zunächst wird das Audiosignal in eine Anzahl n von spektralen Anteilen zerlegt. Das zerlegte Audiosignals wird in einem zweidimensionalen Array mit einer Vielzahl von Feldern gespeichert, mit Frequenz und Zeit als Dimensionen und der Amplitude als jeweils einzutragenden Wert im Feld. Dann werden aus jedem einzelnen Feld und mindestens zwei zu diesem Feld benachbarten Feldern des Arrays Gruppen gebildet, und den einzelnen Gruppen eine Priorität zugeordnet, wobei die Priorität einer Gruppe umso größer gewählt wird, je größer die Amplituden der Gruppenwerte sind und/oder je größer die Amplitudenunterschiede der Werte einer Gruppe sind und/oder je näher die Gruppe an der aktuellen Zeit liegt. Schließlich werden die Gruppen in der Reihenfolge ihrer Priorität an den Empfänger übertragen.